

RECEIVED

Page 1 of 7

FEB 2002

TECH CENTER 1600/2900

KK 2/5/02



1652

RAW SEQUENCE LISTING

DATE: 02/05/2002

PATENT APPLICATION: US/09/478,188B

TIME: 14:14:06

Input Set : A:\896010us.app

Output Set: N:\CRF3\02052002\I478188B.raw

p.5

3 <110> APPLICANT: SHEN, BEN
4 LIU, WEN
5 CHRISTENSON, STEVEN D.
6 STANDAGE, SCOTT
8 <120> TITLE OF INVENTION: GENE CLUSTER FOR PRODUCTION OF THE ENEDIYNE ANTITUMOR
9 ANTIBIOTIC C-1027
11 <130> FILE REFERENCE: 407T-896010US
13 <140> CURRENT APPLICATION NUMBER: 09/478,188B
14 <141> CURRENT FILING DATE: 2000-01-05
16 <150> PRIOR APPLICATION NUMBER: 60/115,434
17 <151> PRIOR FILING DATE: 1999-01-06
19 <160> NUMBER OF SEQ ID NOS: 118
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 42000
25 <212> TYPE: DNA
26 <213> ORGANISM: Artificial Sequence
28 <220> FEATURE:
29 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
30 C-1027 gene cluster sequence
32 <220> FEATURE:
33 <223> OTHER INFORMATION: orf; relative position 658-11
35 <220> FEATURE:
36 <223> OTHER INFORMATION: orf; relative position 1478-930
38 <220> FEATURE:
39 <223> OTHER INFORMATION: orf; relative position 2713-1649
41 <220> FEATURE:
42 <223> OTHER INFORMATION: orf; relative position 3238-2851
44 <220> FEATURE:
45 <223> OTHER INFORMATION: orf; relative position 4971-3442
47 <220> FEATURE:
48 <223> OTHER INFORMATION: orf; relative position 5982-7478
50 <220> FEATURE:
51 <223> OTHER INFORMATION: orf; relative position 9900-7573
53 <220> FEATURE:
54 <223> OTHER INFORMATION: orf; relative position 11349-9982
56 <220> FEATURE:
57 <223> OTHER INFORMATION: orf; relative position 28590-29588
59 <220> FEATURE:
60 <223> OTHER INFORMATION: orf; relative position 29632-31197
62 <220> FEATURE:
63 <223> OTHER INFORMATION: orf; relative position 31280-32590
65 <220> FEATURE:

ENTERED

RAW SEQUENCE LISTING

DATE: 02/05/2002

PATENT APPLICATION: US/09/478,188B

TIME: 14:14:06

Input Set : A:\896010us.app

Output Set: N:\CRF3\02052002\I478188B.raw

```

66 <223> OTHER INFORMATION: orf; relative position 32809-34392
68 <220> FEATURE:
69 <223> OTHER INFORMATION: orf; relative position 35274-34458
71 <220> FEATURE:
72 <223> OTHER INFORMATION: orf; relative position 17924-16653
74 <220> FEATURE:
75 <223> OTHER INFORMATION: orf; relative position 16653-15919
77 <220> FEATURE:
78 <223> OTHER INFORMATION: orf; relative position 15922-14690
80 <220> FEATURE:
81 <223> OTHER INFORMATION: orf; relative position 14643-14212
83 <220> FEATURE:
84 <223> OTHER INFORMATION: orf; relative position 13012-14079
86 <220> FEATURE:
87 <223> OTHER INFORMATION: orf; relative position 12835-11351
89 <220> FEATURE:
90 <223> OTHER INFORMATION: orf; relative position 25564-24986
92 <220> FEATURE:
93 <223> OTHER INFORMATION: orf; relative position 24702-23566
95 <220> FEATURE:
96 <223> OTHER INFORMATION: orf; relative position 22878-21424
98 <220> FEATURE:
99 <223> OTHER INFORMATION: orf; relative position 21407-19926
101 <220> FEATURE:
102 <223> OTHER INFORMATION: orf; relative position 19929-19267
104 <220> FEATURE:
105 <223> OTHER INFORMATION: orf; relative position 19191-18031
107 <220> FEATURE:
108 <223> OTHER INFORMATION: orf; relative position 35938-35516
110 <220> FEATURE:
111 <223> OTHER INFORMATION: orf; relative position 27214-28593
113 <220> FEATURE:
114 <223> OTHER INFORMATION: orf; relative position 25815-27170
116 <220> FEATURE:
117 <223> OTHER INFORMATION: orf; relative position 23546-22875
119 <220> FEATURE:
120 <223> OTHER INFORMATION: orf; relative position 35274-34458
122 <220> FEATURE:
123 <223> OTHER INFORMATION: orf; relative position 37559-38938
125 <220> FEATURE:
126 <223> OTHER INFORMATION: orf; relative position 40986-39367
128 <400> SEQUENCE: 1
129 gtcgactcta gaggatcccg ggtgcggagt aggggttacg gacgaaggag ggggtgcccg 60
130 cgacgcctgc ggcgaagggc ggttccttga gttcgaggcc ggtggcgagg acgacgtggt 120
131 ccgcgtcgag gatctgcgtg tcggggagcg gcccagggcg cagccctcg gtcaggtacg 180
132 gggtagggcc cctgacggtc acctcgaagc agcggtcgtg ggaccgggcg tcgagcgct 240
133 ccccgccgc ttccacaagg acgacgccg gacaggactc ccgtgcggcc tcgaccagtc 300
134 gggcgctcgag gtagtcctg aagatgcggc ggggggcggg gccctgttcg gtgaacttcc 360
135 acgaagccca gcgcggggc cagtcgcgcc ggtcggcctc ctggttggcc cagttgatga 420

```

RAW SEQUENCE LISTING

DATE: 02/05/2002

PATENT APPLICATION: US/09/478,188B

TIME: 14:14:06

Input Set : A:\896010us.app

Output Set: N:\CRF3\02052002\I478188B.raw

```

136 agtcgagcac gtcctcgcgg aacaccgaca tcttgccggc ctggatattg aagacgtggt 480
137 cccagggggtt gccgtcacgg tgataggcga cgcgggccga gcggtaggcg gcgcgccgct 540
138 ccaggaggac gacttcacgc ggtcttctcg cgaaatgaag caggcgtatc gcggtcgccg 600
139 tgccctgccag gccgcgccct acgaccagca ccttggggcg cgcacccgtc atgccatga 660
140 agcctcccc cgtgactcag ggcggcgctg cgcgcgctcc cgtcgggtgc ctcgctgact 720
141 ggaagttccc tgacctggcg tcaactccac tgatccgtaa ggggatcgcg ggagtggata 780
142 cgggtcagggt cgtgcacgat cgtggcacca gacagatcac cacgtcgata ggcaactcgtg 840
143 agccgcgccc ggggctcgac ggggcggggc accggcaggg gcggccgcgt gatcagccgg 900
144 agcctgtccg ggggcgtgcg tgcggggcgt cagctgtcga tgtcgggaac gccagggacg 960
145 tcgatctcgg tgcgggcgta gtggttgaag tagttggtgt agaggttcac ggccacgtgg 1020
146 acgaagacct cggcgagctc ggtgtccgtc catccctgtg ccacggccgc gttccacgag 1080
147 gcgtcagacg cctcgcgccac ttcgcggcg atctccctgg ccacctggac cagtgtctcg 1140
148 agcttcacgt cgtgcgcggg cgtcccccg cgaatcgcca cggctctctc cagcgtgaaa 1200
149 cccgcgacct tcgcgcgacac cgtgtgcgcc gcctggcagt acgcgcacgc gtcgaccgcg 1260
150 cccacggcga gggcgatcgc ctgcgctgtg cgggcgtcga acgttccatg ttcggcgacg 1320
151 gctccgggtga tcgcggcgta ggtttccagg accacggggg aatgggccat tcccccggtg 1380
152 atgttgagca ctcgcccga ccgcttctcc agtcggcgca ggatgtctcc gccggctgcg 1440
153 ggtgcgggtgt cgatggtgtg gacgggaatc cgcggcatgg gaatgcctct cctcgtagt 1500
154 atgggagttc ctcgtccctc cagtctgccc aagcacctcc cccggtgagc tgtcccggcc 1560
155 gccctccggc ccttcttagg caggtcgccc ggtggtgcgg cccaggacg tcacctcgcc 1620
156 gcaccacgg gcgcccag gggcgaggtc agaggccgag cacctcctcg gccaggcgcg 1680
157 tgccccgaac acgggcctcg atcttggcga aggccaggtc gcgtgtggtg gaggtgtcgt 1740
158 cggcgaacgg ggagaagccg cagtctgcgc aggttcccag ttgctcgacg ggcagtgcg 1800
159 gggcgcgag caggatgcgg tcgcgtacct gctcgggggt ctcgaccact ggtcgtatcg 1860
160 ggtcgggtcac cccgaggaag acgcgggcgg cagggggcag gtggtcacgg acgatgctca 1920
161 ggacccgctc ggggtccgct tcgcgggcca gttcgagata gaagttgcc gccttgagct 1980
162 ggaagagctt gggcagcagt tcggcgtagt cgatgtcgag gctgtgcgtg gagtccgtgt 2040
163 cgcgcgggg gcaggtgtgt acgccgatgc gggcggtttc ctcggcgctg aagcgcccca 2100
164 ggacttcggt gttgagggcg atgaagtcgt cgaggacgcc gccgctgggg tcgagcttga 2160
165 gggacagccg cccctcgggt aagtcgagct ggaccacgtg tgcccccgcg tccaggcagc 2220
166 ctcggatgtc ggttcgggcc tcgtcggcga ggtcgcgcag gaactgctcg cgggggtagc 2280
167 cctcgatggg agtggcgggg tagaggaggc tgagggcgga ggtgcgatg accgcctgct 2340
168 tcagggggcg gtcctgtagc tgccgtgcgg cgcgcagata ggtttcggcc cgcacctggt 2400
169 agcggaaggg ccttgggtg atgctgggga gctgcgggt gtgccgtct gcgaagggga 2460
170 tgacagcgcc gtcgggcgag aggggtgcga ggcgggtcac ggggtagggt gcgaagctcg 2520
171 gcttggactg ttcaccgtcc acgaggacgg ggtgcggac tcgttccagt cgtgtcaggg 2580
172 tgtccgcgac ggcctgttcc tgcgttttgg ccagggtccgt ggcgtccagg gttccctggg 2640
173 catgcgcggc aagggcgtgc aggagtgcg cggagcgcg aaggctgcc atcggtcag 2700
174 tggcgatggt catggccgaa gtagtgggaa gaggtgggt ttcgaaaccac cgcaaagctt 2760
175 tgattgccgc tttttcaggg gaagttgat cgaagtcgcc gagcggcgga acgtgctgat 2820
176 gtatgggggg cgggaggagc ctgcgggggt ctaggagccg gtcgcggcca cggtgaggga 2880
177 ggtgcccgag tgggagcggg ggtcttttc gccgacgcgg ttgggctcga tggtgcgggg 2940
178 gtcgacggcc tctccggggg caccttgccg gtacagccct tcggggtcgg agtcccggtc 3000
179 atgggggagc aggaagaaga cccggcgccg gtacagaccg ctgtccgggt ccgcttcggc 3060
180 gtcggccccg agttcgatgt agccgatcat gcggccgtcg cgggcgtagc gcggcttgtt 3120
181 cttgcgcggg ggggtcttgt ccagggcctg gcggacgtag tcgagtccct cgggatcttc 3180
182 gagccacacg accttcgcct cgtgaacgag atcgctgtcg gtcagtagcg agctcatggc 3240
183 ggcgacctct ccttcgtcgg cgtgcaccgg gtgggggaagc ggtgcctcg tgatgtgtgt 3300
184 tcgtctgcgg cggtgggccg cagtgtgtcg gaccgcccgt ggtgcgggtt ctcggccaaa 3360

```

RAW SEQUENCE LISTING

DATE: 02/05/2002

PATENT APPLICATION: US/09/478,188B

TIME: 14:14:06

Input Set : A:\896010us.app

Output Set: N:\CRF3\02052002\I478188B.raw

```

185 gcacgggcag gtacgtcctg gggcactcac atcgtagatg gggcccgctt ccgcagggca 3420
186 gtgcctccgg tcggaggacg ttcattcgtc ggctgccaga gcgaggttgg ggtagaactt 3480
187 ccggccgttg gatttgatca tgtcggcagg tgaggcgagg ccacttcctt ggccgacccg 3540
188 ggtggcgaag gcacgggcgg tcccggggcg gatgccttca ctgtgtgcgc accaggtgct 3600
189 gtaggacgtg tagagaaggc cctgttcgac gcgtagctcg ctgttctcgg ggtcgtggag 3660
190 gcagcactcg gcgaggaagc ggccgatgtg gtcctcgggtg ttcgcgtatg cgctggtggc 3720
191 gatcgcgacc cgtcgggggc cggcgagtgt gtcgcgggtg gcgaggtagc ggccggcccc 3780
192 ttcggtgagc cagtgcagga tcccggggcc ctgcctctgg acgagttcga cagccaggtt 3840
193 gtcgatcttg cgttcgtcgg ggacgatccg ttcgaagggc aggaggcgga tgcggcgcca 3900
194 gaaggcgaag ccgccggttg agacctcggg gcggtggttg ccagcagcc acagcttgtg 3960
195 cgtgggtgtg aaggagaaat agtctgcgg catgcggcgg gccttgatct tgtcaccgcc 4020
196 ggtcagcagg cggacgcgcg cctcgtcgaa gcggtcgttg ggcttgagct cgctgcacac 4080
197 gatgaggcgg cggccgtgga gttcgggtgag ctccggtggag tgttcggagt atgcgccacg 4140
198 gtccatgagg aaaccgcggc gggctgcgtc ggcgtagctg ccgagaatct ggatcatcac 4200
199 gtcgaggaga acggttttgc cgttctttcc ctggccgttg agaaagggca gcacctgcgc 4260
200 cccgacgtca ccggtgatgg agtagccgag aaggaggttg aggaagtcga tcactctccc 4320
201 cccttcggcg tcaactgccg aggtgtcttc gaggaaacgg tgccagcggg ggggtgggat 4380
202 gtccctgggg gaggcgctgg tggcgcgga gtggaagtcc cgggtggggt cgggcttgcg 4440
203 catacggccg ttgcggaggt cgaccactcc gtcaggggtg cacagggcgt aggggtctcc 4500
204 gtcgaggggt tcgggatcga gggagaggtc gggagaggcc tttgcctggg tgaggagcgc 4560
205 cttcataccg gtcgtcgaca ggggtcggcg tttgtggtgg tgcagttccc ggtcggtgaa 4620
206 cagcccgcg ggatcgctgc cgggcctcct ctccgccatc tctccggcag cccacaggcc 4680
207 agctttctcg cctccggccc gcttccaccg gtagccgtcc caggagtacc agccaggcc 4740
208 ctccacgtgc cggaaactgg cagcgtagag acggacgaag agcttgccgt tgcgcggtc 4800
209 ggtcaggctg gcgggaatct cggccgcctc ccaggcggtc gcggcgacgg gggcctcggg 4860
210 agcggcctgg acagggagga gcggcgctgg ggcgggggtg gtttcgaggg ccagcatctg 4920
211 ctgagcggcg gcagttgcgt caaagcgagg gccctcggcg ctgctgctca tggacgtcct 4980
212 tcgagatgga gcggtcgggc ggtccccgct gcgggaacgg catgaatgat cttcccggtg 5040
213 cggacagagt gccaggggca gcgcatgtgc ggggggacaa cggcccgttt cggacgaggg 5100
214 ccggccgacg gggggaagca ggggccggca accgggtggc ggggcggcgt gagcgaggc 5160
215 acgagcggcc cggtagcggg ggaagggtc gtctctccgt ggggcggcac gttgtggtcc 5220
216 tcgtccgtca gcttgctctt ggcttcagcc tcctgacccc caataaggcg aaagctgctg 5280
217 gtcaagcatc tttcgtgaca ctccgagagg gactgaaggg actgtctttc ggaatgagtg 5340
218 tagggggttg tcgggtgggg accgcgcctc gactccccgg cggacgggat ctgttcggtc 5400
219 ggtcccttgg gtccctcccc ggatcgcggc agggacccaa ggggcgggtg cggcgggcgg 5460
220 tcggtgaggg gccccggttg agggactgag ggtctgtatg gacgataag aggtctgaa 5520
221 ggggcggaga gagtttcggt ccctgcgttg agtccctggt catcaccgca ggtcagaggg 5580
222 gttttgaggg gtgaaaaagg gactgaagg actcaacttc ccattatga gctgagtaga 5640
223 agaaagcagt atgacgatat cggcgccctac atacgcgcgc gtacatagtg agcttataat 5700
224 cgggaagtgg agtcccttca gtcccttttc gtgggtcgt atccctctg actgcgttga 5760
225 ccgtcgccgc tccgcgcagg gaccgaagag ggaccaagtc cctgcgcggg gcgggcgacg 5820
226 gtaatcgtgc agtgccttcc ccccgcttcc ccacagcgag tcgtcgctcc cctgtgaggc 5880
227 cggagagggg cctagaacct ctcaggggcc gttctgtggc cctctgggcc tctcctggc 5940
228 catttaccct atgggggcgc ttgggggcgt caggagggt tgtgagggt ctgccgggaa 6000
229 gtggcggatt gcgcatggca ggagatgcc cgacagcggc cgggaatcga cgatgtcccc 6060
230 cgaccctat ccagcgtccg ctgatcctca ggaggcagac cttgcaggct ccagaagcga 6120
231 agaacggccg gtccccggag cagccgcagg aagagcggat cgtcctggac gtatggctgg 6180
232 cgaactacct gttccccacc tatgacgggc gtgacttccct cgctccgctg cgcgagcggg 6240
233 cggcggagtt cgagcgcgcc caccctcgat accgggtcga catcaacggc cagacttct 6300

```

RAW SEQUENCE LISTING

DATE: 02/05/2002

PATENT APPLICATION: US/09/478,188B

TIME: 14:14:06

Input Set : A:\896010us.app

Output Set: N:\CRF3\02052002\I478188B.raw

```

234 ggaccatccc cgagaagggtg gcgcgcgcga ccgcggaggg caggcctccg cacatagcgg 6360
235 gctactacgc caccgacagc cagttggcgc gggacgcgcg caggcccgcg gggaagccgg 6420
236 tcttcacctc ggtggaggcc gcgttggccg gccggacgga gatactggga cacccggtgg 6480
237 tgggtggagga cctcgacccc gtggtgcgcg actcctactc gttcgggggc gagttggtgt 6540
238 cgctgccgct cacggtcacc accatgctct gctacgcca ctcctccctc ctgcgcgcgc 6600
239 ccggtgttcc ggagttgccc cgtacctggg atgaggtcga agcagcctgc caggcggtgg 6660
240 ccagcgtcga cggggggccc ggtcacgga tcacctgggc caacgacggc tgggttttcc 6720
241 agcaggccgt cgcccttcag aacggggtgc tgaccgatca ggacaacggc cgtccgggt 6780
242 ccgccacgac ggtgcgagtc acatcggagc agatgctgga ctgggtccgc tgggtggacgc 6840
243 acctccatga gcgcggccat tacctctaca cgggcggggc ctcgactgg ggccggggcgt 6900
244 tcgaggcttt cgtccagcag aaggtcgcac tcaccttcga ctggtccaag gccgcccggg 6960
245 aactcatcca ggccggtgca caggccggtt tcgaggtcgc ggtgttcccc ttgccagga 7020
246 acgcgaaggc cccggtagcg ggccagcccg tctcgggaga ctccctgtgg ctggccgcgg 7080
247 gactcgacga gaccacgacg gacgggctgc tcgctctcac ccagtacctg atcagcccgg 7140
248 ccaacgccgc ggactggcac cgcaccaacg gtttcgtacc ggtgaccggc gcggccgggg 7200
249 aactgctgga agcgacaggc tggttcgacc gccggccgca gcaacgggtg gccggggagc 7260
250 agttgaaggc gtccgaccgg tcaccggcgg cgtcgcggcg gctgctcggc gacttcgcgg 7320
251 ccgtcaacga ggtcatcacc gcagcgatgg acgatgtcct gcgcagtgga gccgaccccg 7380
252 cgaaggcctt cgccgaagcc ggctggcccg cccagcaact gctcgatgcc tacaacgccc 7440
253 ggaaccgctc cggatccggg accccctccg ccgtctgaga tccggtaccg gggcacaggg 7500
254 gcgccgccgc ccgctttccc ggccggggcac tggccggggg acatgctctc ccgcccccg 7560
255 caggacgtag ggtcaacccg cctgcgcctt caggtggcgg cgcagatact caccggtcag 7620
256 ggaggaatcc gcggcgagca ggtccttcgg tgtgccgggt aagacgatct cgcgccttc 7680
257 cgtcccccg tcgggaccca ggtcgatgac ccagtcggcc tgctgcacca catcgaggtt 7740
258 gtgctcgatg accacgacgg tgttcccgcc ctcgacgagc ccgtccagga gcttcagcag 7800
259 ggtgtcaacg tccgacatgt gcagcccggg ggtgggctcg tccaggacat agaccgtgcc 7860
260 cgtgcggtgc agctggtcgg caagtttgat ccgctgcagt tcaccgcccg agaggctgga 7920
261 aagcggctgg cccaggctga ggtacccaag accgacgtcg acgagagcgc gcagtttcgg 7980
262 cagcagggcc ttctcgggtg agaactcgac ggctcgtcg gcgggcagct ccaggacgtc 8040
263 cgcgatcgac ttcccgcgaa gctggtgctc caggacctcg ggcttgaagc ggcccccctc 8100
264 acagacaccg cagtgcgtgg tcaccggatc catgaaggcc agctcgtgta tgatgacccc 8160
265 gcggccctgg cactcctcgc acgaccctt ggagttgaag ctgaacagcg aggcgttcgc 8220
266 gccggtctcc ttcggaaca gcttgcgcag cgggtccatc aggccgaggt aggagaccgg 8280
267 tgtggagcgc gacgaggcgg cgatcgcgga ctggtcgaca aagaccgcgt cggggtgcgc 8340
268 ctccatgaat gcccgggaga tcaggctgct cttgccggaa cccgccccc cggtcaccgc 8400
269 ggtcagcaca ccggtgggca cggccacgga gacctgcttc aggttgtgga gatccgcgtt 8460
270 ctccacggtc agtcccccg tggcgggcg gacctctcc ttacgcggg ccccccgcc 8520
271 cagagcctcc ccggtccggg tcttcgcctt ccgcagcttc gcgaaggacc cctcgaacac 8580
272 gatctcgccc ccgtgcactc ccgccccggg accgacatcg acgatgtggt cggcgatctc 8640
273 gatcacatcg ggtcgtgct cgacgaccag cacggtgttc cccttgtcg gcagcgcgcg 8700
274 cagcagtcg ttgagcgccc ccacgtcgcg cgggtgcagg ccgatgctgg gctcgtcgaa 8760
275 gatgtacgtg agcccggcca gaccactgcc gaggtggcgc accatcttca gccgctgccc 8820
276 ctgccccccc gagaggtcgg ccgtgggcct gtccagggtc aggtagccga gcccgatgga 8880
277 cagcatccgc tccagggccg tcgcgcgggc tttcgcgaga ggggcagcgg ccggtccgt 8940
278 gacgccggcg agcacctccg tgaggtcgcg gacctccatg ctcgagtagt cggcgatgtt 9000
279 cttgccgtcg atccggacgt cgagcgcggc ggcgttgagc cgcgcgcccc ggcaggaggg 9060
280 acagactccg tcggtgacga aacgttcgat gacctcgcgc ttgcggtcgc tcagcgcgct 9120
281 gaggtcgcgc ttgaggttga gccgctcgaa ccggtcggcc aaccctcgt agttcgtctg 9180
282 gaactcgggt ctcttggtct tcagcgtcac cttcccgcgg gtgccgcgca gcagcgtgtc 9240

```

Use of n and/or Xaa has been detected in the Sequence Listing.
 Review the Sequence Listing to insure a corresponding
 explanation is presented in the <220> to <223> fields of
 each sequence using n or Xaa.

VERIFICATION SUMMARY

DATE: 02/05/2002

PATENT APPLICATION: US/09/478,188B

TIME: 14:14:07

Input Set : A:\896010us.app

Output Set: N:\CRF3\02052002\I478188B.raw

L:2477 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:104

L:2494 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:105